DOT/FAA/AR-95/58

Office of Aviation Research
Washington, D.C. 20591

Bulk Explosives Detection System Developmental Test and Evaluation: General Protocol for Screening Baggage and Electronics

Ronald A. Krauss
A. Kenneth Novakoff

Aviation Security
Research and Development Division
FAA Technical Center
Atlantic City International Airport, NJ 08405

September 1995

General Test Plan

U.S. Department of Transportation
Federal Aviation Administration

NOTICE

Technical Report Documentation Page

1. Report No. DOT/FAA/AR-95/58	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle BULK EXPLOSIVES DETECTION SYSTEM DEVELOPMENTAL TEST AND EVALUATION: GENERAL PROTOCOL FOR SCREENING BAGGAGE AND ELECTRONICS.		5. Report Date September 1995 6. Performing Organization Code AAR-520
7. Author(s) Ronald A. Krauss, A. Kenneth Novakoff		8. Performing Organization Report No.
9. Performing Organization Name and Address Aviation Security Research and Development Division Office of Aviation Research Federal Aviation Administration Technical Center Atlantic City International Airport, NJ 08405		10. Work Unit No. 11. Contract or Grant No.
12. Sponsoring Agency Name and Address U.S. Department of Transportation Federal Aviation Administration Office of Civil Aviation Security 800 Independence Ave., S.W. Washington, DC 20591		13. Type of Report and Period Covered General Test Plan 14. Sponsoring Agency Code ACP-400

15. Supplementary Notes: Assistance in report preparation provided under contract #DTFA03-92-C-00027 by: DCS Inc. (D. Lamb), 2511 Fire Road, Suite A-5, Absecon, NJ 08201

16. Abstract
Development of devices and systems for detection of explosives in passenger luggage is supported by test and evaluation. This document provides a framework for the structured implementation of developmental test and evaluation for the bulk explosives detection program. Planning considerations, general test objectives, evaluation techniques and test management controls are discussed. Guidelines are presented for preparing detailed test and evaluation plans and summary reports for each test and evaluation effort. Standard forms to be used in testing are provided as an appendix.

17. Key Words	18. Distribution Statement
developmental test and evaluation bulk explosives detection aviation security	This document is available to the public through the National Technical Information Service, Springfield, Virginia 22161.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	39	

Form DOT F 1700.7 OP (3-89) Reproduction of completed page authorized

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS AND SYMBOLS

COTR	Contracting Officer's Technical Representative
DPR	DT&E Problem Report
DT&E	Developmental Test and Evaluation
EDD	Explosive Detection Device
EDS	Explosive Detection System
FAA	Federal Aviation Administration
IT&E	Independent Test and Evaluation
R&D	Research and Development
ROC	Receiver Operating Characteristic
SOP	Standard Operating Procedure
p_d	probability of detection
p_{fa}	probability of false alarm

1. INTRODUCTION.

1.1 Test Management Personnel Definitions.

There are various categories of personnel who will be involved in the management of bulk Explosives Detection System (EDS) Developmental Test and Evaluation (DT&E) which are defined below. Individual roles and responsibilities of all personnel involved in the conduct of DT&E are delineated later in Paragraph 4.6.

Program Lead. The Program Lead is the Bulk Explosives Detection Program Lead and is responsible for providing overall direction for bulk EDS research and development (which includes DT&E).

Project Manager. The Project Manager may be the Contracting Officer's Technical Representative (COTR) for the project or an individual designated by the Program Lead to be the technical lead for the project.

Test Manager. The Test Manager will be designated by the Program Lead to oversee test conduct, and may be the same individual as the Project Manager.

Test Director. The Test Director is independent of the Bulk Explosives Detection Program and is responsible for EDS independent, operational and certification testing. The Test Director will generally support DT&E as requested.

1.2 Background.

Section 107 of the Aviation Security Improvement Act of 1990 (Public Law 101-604) requires the Federal Aviation Administration (FAA) to establish and carry out a program to accelerate and expand the research, development, and implementation of technologies and procedures to counteract terrorist acts against civil aviation. As part of this effort, the FAA has fostered the development of explosives detection devices (EDD's) and explosives detection systems (EDS's) by way of contracts, grants, and cooperative research and development agreements with industry, colleges, universities, and other research institutions.

In support of this mission, the FAA is developing two classes of explosives detection technologies: bulk detection and trace/vapor detection. Bulk EDD's and EDS's detect quantities of explosives based on their elemental and structural composition. Trace/vapor EDD's detect trace amounts of explosive particles and/or vapors emitted from the explosives or on contaminated surfaces based on their chemical and physical properties.

The FAA will periodically assess the performance of systems under development by conducting system test and evaluation. This testing, defined herein as Developmental Test and Evaluation (DT&E), will be performed by the developer of the EDS (the "vendor") and/or the FAA and its technical representatives. Testing may be conducted during or between development stages, i.e. laboratory prototype, engineering prototype, commercial prototype, etc., depending on specific test objectives or contractual requirements. DT&E may also include studies which may not directly assess performance but which provide data required in the development process.

DT&E may also be conducted on commercially available devices. Testing of this nature will generally be performed in cooperation with the EDS vendor and will be done to enhance the overall goals of the bulk explosive detection research and development (R&D) program and to provide data to the vendor which would not otherwise be readily obtainable (because actual explosives are used).

DT&E will be performed under the direction of the bulk explosives detection Program Lead. DT&E will either be distinguished from or augment Independent Test and Evaluation (IT&E) performed by the Aviation Security Research and Development Test Director. IT&E is generally an independent assessment of system performance requested by the Assistant Administrator for Civil Aviation Security, and may be performed in operational environments (i.e. airports) using a set of test objects based on the EDS Certification Criteria (see ref. D). While using the certification criteria as a guide, DT&E objectives and procedures will be formulated to achieve R&D goals which may be broader in scope than limited performance assessments. For example, DT&E may consider threats which are not included in the certification criteria, or may perform a study of EDS performance for a variety of EDS internally variable parameters. In those cases when it is desired to compare performance of different systems, however, a set of procedures will be formulated for performing baseline performance assessments which will provide statistically meaningful comparisons.

1.3 Purpose.

Methods of conducting test and evaluation of vapor/trace detection systems differ markedly from those for bulk detection systems. Therefore, system DT&E will generally be controlled by independent protocols. The purpose of this plan is to provide the framework to ensure a structured implementation of the FAA's *bulk* EDD/EDS DT&E effort. This plan will define the activities which will be performed during DT&E of bulk EDD's and EDS's and provide guidance for the development of detailed test plans for specific DT&E efforts.

Note: "EDS" refers to a device that has been designed to meet the EDS certification criteria (Reference D) whereas "EDD" refers to a system which has been designed to meet part of the performance requirements contained in the criteria, or other requirements. For simplicity, both will be referred to as EDS's throughout this document.

1.4 Scope.

This plan addresses the entire DT&E process from planning considerations through system evaluation methodology. In particular, this plan includes: test program objectives, planning considerations, general test and evaluation approach, training requirements, safety requirements, test schedule, documentation requirements, test control, and roles and responsibilities.

Note: This plan, while not excluding the possibility of airport testing or data collection, does not include special considerations for such scenarios. A separate document may be generated for airport data collection and developmental test and evaluation. In addition, testing conducted at vendor facilities will require special provisions not addressed in this plan.

1.5 Applicability.

This plan contains requirements for the conduct of bulk EDS DT&E. However, since DT&E will be performed on systems at various stages of development, some sections of this plan may not be applicable to all testing efforts. For each DT&E effort, the Project Manager will be responsible for determining test objectives. The cognizant Program Lead will determine the applicability of this plan to the test objectives, thus ensuring that minimum test requirements are met. The Project Manager will then be responsible for ensuring that the relevant DT&E effort complies with this plan as applicable. The test will be conducted by a designated Test Manager.

All FAA-conducted DT&E will be conducted in accordance with certain minimum requirements dictated by this plan. These minimum requirements are:

A. A Detailed Bulk EDS DT&E Plan will be developed in accordance with Paragraph 4.4.1.
B. An FAA laboratory notebook will be required for manual recording of test data, observations, calculations, etc., in accordance with Paragraph 4.4.2.
C. All testing will be conducted in accordance with safety plans and procedures contained in References E through H.
D. Hazardous emissions testing will be conducted in accordance with Paragraph 3.4.2.5.
E. A Bulk EDS DT&E Report will be written in accordance with Paragraph 4.4.3.

2. REFERENCE DOCUMENTS.

A. FAA Order 1810.4B, FAA NAS Test and Evaluation Policy, October 22, 1992.

B. FAA-STD-024b, Preparation of Test and Evaluation Documentation, August 22, 1994.

C. DOT/FAA/CT-93/54, Management Plan for Explosives Detection System Certification Testing, September 24, 1993.

D. FR 47804, Criteria for Certification of Explosives Detection Systems As Defined Under 14 CFR 108.20 (U), September 10, 1993.

E. DOT/FAA/CT-ACA001A94/05, Safety and Health Plan for Bulk Explosive Detection Systems Testing, June 1994.

F. DOT/FAA/CT-ACA001A94/08, Standard Operating Procedures: Handling and Storage of Explosives, June 1994.

G. DOT/FAA/CT-ACA001A94/03, Standard Operating Procedures: Threat Article Preparation, Transportation, Storage and Handling, June 1994.

H. DOT/FAA/CT-ACA001A94/02, Standard Operating Procedures: Receipt and Transportation of Explosives, June 1994.

I. Classification Guide for FAA Explosives Detection Systems Information and Data, November 21, 1990 .

J. Baseline Test Protocol (to be developed).

K. Test Article Selection Plan (to be developed).

3. TEST PROGRAM DESCRIPTION.

This section defines the approach and provides the framework to be followed throughout DT&E, and includes a description of DT&E objectives, planning considerations, and the general approach to test conduct and data evaluation.

3.1 Test Program Objectives.

Bulk EDS DT&E has three primary objectives: specification validation, system/technology comparisons, and strengths and weaknesses determination. These three objectives are discussed in the following paragraphs.

3.1.1 Specification Validation.

DT&E will be conducted on systems at various stages of development to evaluate progress and to assess how well contractual specifications or acceptance criteria are met, such as minimum explosive detection or throughput rate. DT&E may also be conducted independent of contractual requirements to assist in the engineering design and development process. In this way, DT&E will be instrumental in minimizing the risks associated with proceeding toward the next project milestone.

3.1.2 System Comparisons (Baseline Testing).

Baseline testing will be conducted in order to provide data to compare performance of different systems and technologies. This comparison analysis will provide direction for future bulk program efforts and will also provide data for combined technology program efforts. In order for system comparisons to be valid, baseline testing must be conducted using standard procedures and test articles. Performance comparisons may include detection performance with respect to certain explosive types, masses, and configurations; false alarm performance; and throughput performance. As much as possible, baseline testing will be conducted for all systems which are undergoing DT&E at the FAA Technical Center.

3.1.3 Strengths and Weaknesses.

DT&E will be conducted to evaluate the detection performance limits of the system under test. This capability assessment will be accomplished by evaluating the EDS using a test article set that is specifically constructed to test the system's strengths and weaknesses. This test article set will include explosive types, quantities, and orientations which are suspected to approach the performance limits of the EDS under test.

Testing will also be conducted as appropriate during system development to evaluate limitations inherent to the system/technology. Potential system vulnerabilities may be evaluated by conducting tests utilizing interference samples and operating the system under potentially adverse conditions (e.g. in the presence of electromagnetic interference, high temperature, humidity, etc.). The interference samples may consist of probable false alarm items, clutter, and known countermeasure techniques and will vary depending on the technology employed by the EDS. The interference samples will be selected based upon

theoretical analysis of the vulnerabilities of the system/technology, and upon previous T&E experience.

Note: All system performance data will be reviewed for classification in accordance with Reference J.

3.2 Planning Considerations.

The following paragraphs describe the planning activities which are required for the successful execution of EDS DT&E. These planning considerations should be used as guidance in development of the Detailed Bulk EDS DT&E Test Plan.

3.2.1 Test Objective Determination.

The determination and prioritization of DT&E objectives must be made during the planning stage. The DT&E objectives will dictate the required evaluation criteria, test data, data confidence levels, and test resources (e.g. personnel, equipment, facility, time, and test articles).

3.2.2 Data Requirements.

The required data (e.g. raw data, system alarm/no alarm, x-ray images) must be identified during the planning stage. Provisions must be made for recording data, which could be in the form of manual recording of performance data in a laboratory notebook, recording of system data to computer disk or magnetic tape, recording of optical images, etc.

3.2.3 Data Confidence Intervals.

The minimum allowable statistical error and confidence interval must be considered during the planning stage. Determination of the minimum desirable confidence interval will dictate the required number of trials (and hence test articles) and therefore will affect the duration of the test and other factors.

The data confidence intervals are dependent on the chosen confidence level, number of independent trials, and measured detection or false alarm rate. Figure 1 and Figure 2 are two-sided and one-sided confidence interval charts to be used during the test planning stage. The charts plot number of trials vs. confidence interval for six different confidence levels, ranging from 60% to 99%. A discussion of confidence intervals and an explanation of confidence interval calculations to be performed during data analysis (for one-sided and two-sided tests), is provided in Paragraph 3.4.3.1.

The Project Manager will be responsible for determining the confidence level, the number of trials, and the acceptable confidence interval for each test.

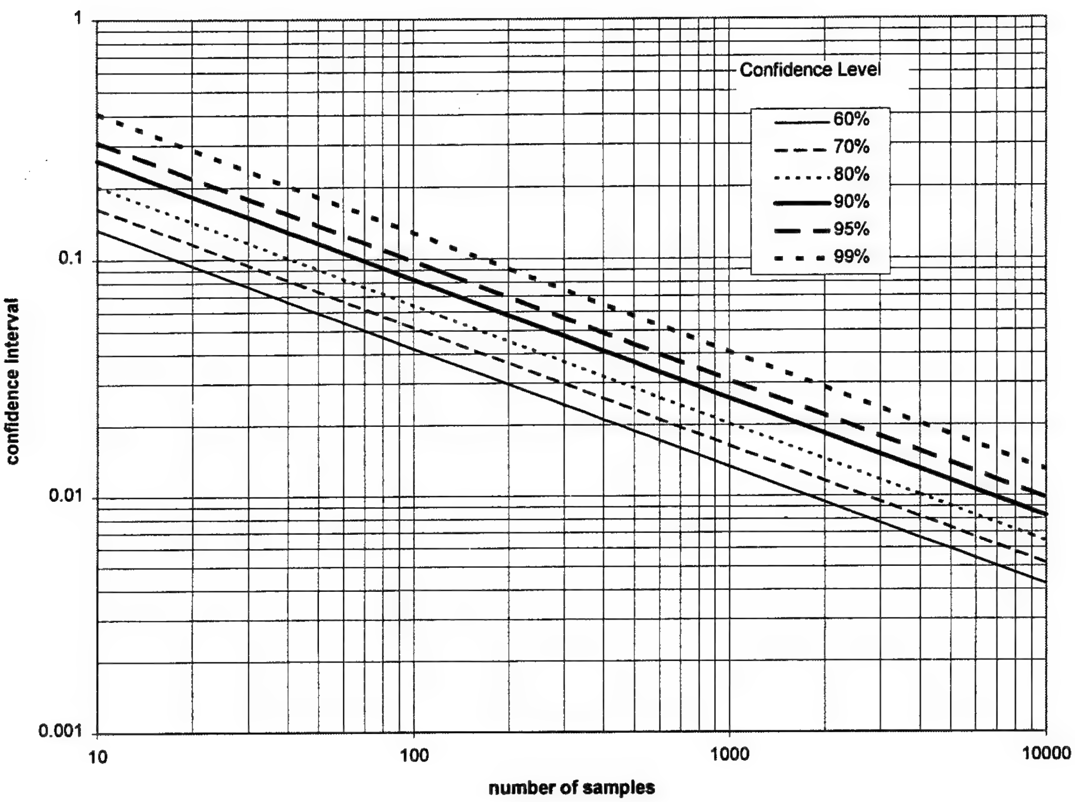


Figure 1. Two-Sided Confidence Interval Chart. For a given confidence level, one can determine the maximum required number of independent trials corresponding to a desired confidence interval (statistical error), or one can determine the maximum confidence interval for a given number of independent trials. A probability of 0.5 is considered "worst case" since a lower or higher probability will result in either fewer required trials or a smaller confidence interval. (See text for explanation of two-sided vs. one-sided confidence intervals.)

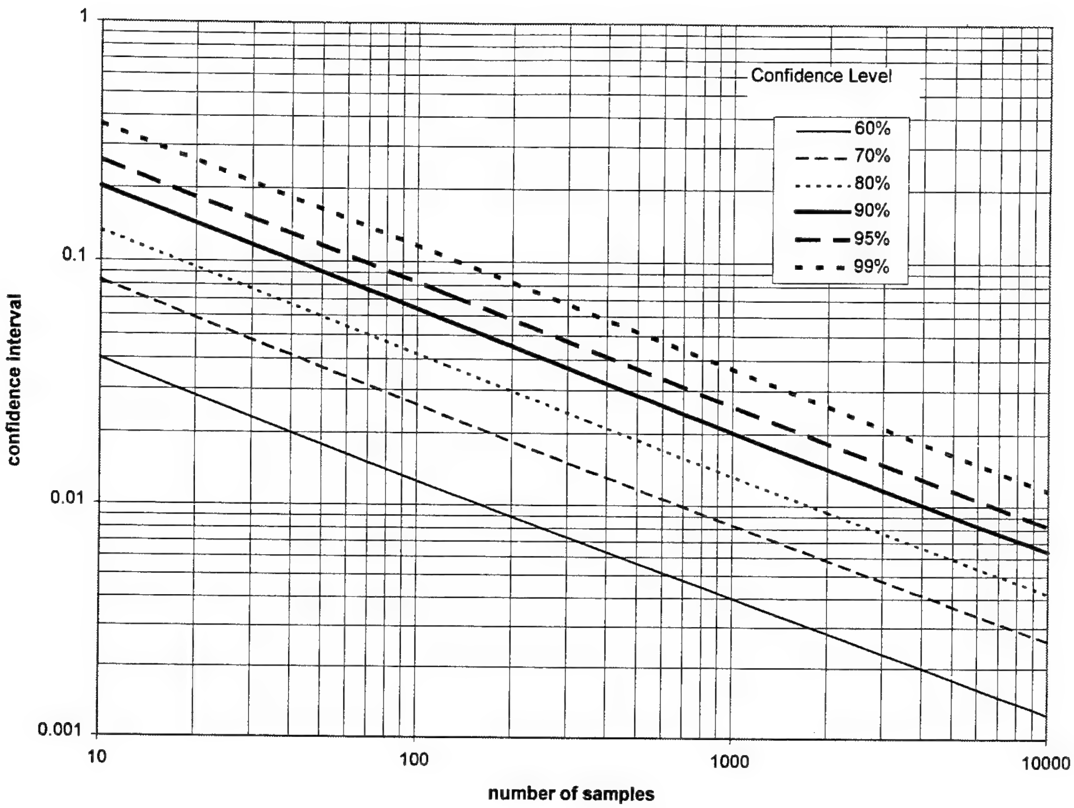


Figure 2. One-Sided Confidence Interval Chart. This method of determining sample sizes or confidence intervals may be more appropriate for certain applications in bulk DT&E, such as determinations of probability of detection or false alarm.

3.2.4 <u>Personnel Requirements.</u>

Test personnel critical to the execution of the DT&E effort must be identified during the planning stage. This includes identifying any necessary technical support services (e.g. FAA contractor, EDS vendor), and funding to obtain those services. Proper planning will help to prevent personnel schedule conflicts which may result in test delays.

3.2.5 <u>Equipment Requirements.</u>

All equipment required for the execution of DT&E (including EDS's, test and analysis tools, conveyor systems, etc.) must be identified during the planning stage.

3.2.6 <u>Facility Requirements.</u>

Facility requirements such as physical, electrical, and cooling requirements must be identified during the planning stage. Proper planning will obviate the need to postpone testing due to possible facility modification lead times. Any other system-unique requirements, such as facility licensing requirements, must also be considered. In addition, all FAA Aviation Security laboratories required for DT&E must be identified and scheduled as early as possible.

3.2.7 <u>Time Requirements.</u>

The amount of time required to conduct DT&E must be estimated during the planning stage. Where possible, time requirements for each phase of testing should be estimated to allow for dynamic allocation of resources.

3.2.8 <u>Test Article Requirements/Availability.</u>

The number and types of test articles required for a specific DT&E effort are dependent on predetermined test objectives, confidence intervals or sample size. Therefore, in order to ensure that DT&E objectives are satisfied, it is necessary to consider the availability of test articles during the planning stage. The required set of test articles will be defined by the Project Manager. If obtaining data for system comparisons, the minimum set of test articles is specified in the baseline test protocol. Availability of baggage, electrical items, and explosives must be determined as soon as possible to allow for reevaluation of test objectives if necessary. Primary consideration must be given to the logistics involved with transportation of test articles and explosives, especially to test sites outside of the FAA Technical Center.

3.2.9 <u>Security Requirements</u>

Provisions for handling of security-sensitive or classified data need to be identified. In the case of vendor facility testing, a review of the vendor's security clearances, procedures and facilities should be conducted.

3.3 General Test Approach.

The general test approach includes standard practices for the conduct of all EDS developmental testing. Specific areas addressed in this section include a description of the test environment, test article set, configuration management practices, system calibration and threshold verification, pretest activities, data collection, and problem reporting.

3.3.1 Test Environment.

The test environment will be created to ensure that there are no factors, such as electromagnetic interference or improper grounding, which will adversely affect the outcome of a test. (Unless the test objective is to determine environmental effects on EDS performance.) Test operating conditions (e.g. temperature, humidity) will be kept within the limits specified in the contract and/or vendor specification, as much as possible. The test environment will be monitored throughout testing, as deemed appropriate by the Test Manager.

3.3.2 Test Article Set.

A DT&E test article set will be established and will consist of a standard test article set, and/or a custom test article set as applicable to test objectives. These test article sets will be developed in accordance with the DT&E Test Article Selection Plan and the standard test article set will be specified in the Baseline Test Protocol (see Paragraph 4.4.4). An overview of the two test article sets and test article identification methods is presented in the following paragraphs.

3.3.2.1 Standard Test Article Set. The Program Lead and Test Director will develop a standard set of test articles for bulk EDS DT&E. The planned utility of the standard set is to provide for a baseline test for all EDS's undergoing DT&E at the FAA Technical Center. Usage of the standard test articles and standard test procedures will ensure that performance data are collected for the same test articles and have sufficient statistical confidence to allow for comparison of systems or technologies.

The standard set will be divided into two categories: baggage (checked and carry-on) and electrical items. These categories will be divided into two subsets: the threat subset and the non-threat subset. The threat subset will contain explosives and will be used to measure EDS detection performance. The non-threat subset will not contain explosives and will be used to measure EDS false alarm and throughput performance.

Both the threat and non-threat test articles will consist of preselected baggage and electrical items, which will be stored at the FAA Technical Center. The DT&E Test Article Selection Plan will provide baggage and electrical item selection parameters for both the threat and non-threat subsets. The DT&E Test Article Selection Plan will also specify a standard set of explosive parameters (composition, quantity, orientation, and position in the test article), which will be used to assemble the threat articles prior to each DT&E effort.

Note: All explosives will be returned to the FAA's explosive storage magazines at the conclusion of each day of DT&E. It may be required to remove explosives from test articles.

3.3.2.2 Custom Test Article Set. Since EDS DT&E can have a multitude of objectives (e.g. parametric testing, countermeasure evaluation, environmental effects, etc.), and is conducted on systems at various stages of development and/or at remote locations, it will not always be appropriate or feasible to conduct DT&E utilizing the standard test article set. During the planning stage, the Project Manager will determine if a custom test article set will be required. If required, the Project Manager will develop a test article set that can be utilized to meet the defined test objectives. Special provisions for obtaining test articles will have to be made well in advance of test conduct in those cases when they are not from FAA stocks, e.g. from local law enforcement agencies. The test article selection process will be described in the Detailed Bulk EDS DT&E Plan.

3.3.2.3 Test Article Identification. Where practical, each test article will be uniquely identified with a human and/or machine readable label (barcode). Relevant information pertaining to the test articles will be separately recorded. This information may include a description of the explosive (type, mass, configuration) and its location in the test article, test article characterization data, and a list of test article contents that may influence test results. Other means may be used to document the test article contents and explosive positions within the test article including photographs, videos, and collection of EDS images.

3.3.3 Configuration Management.

If applicable to the EDS, system configuration will be established during a configuration audit prior to the start of DT&E. At this time, designated test team personnel will identify and record version and serial numbers of applicable hardware and software, including diagnostic and maintenance tools. Data collected during the configuration audit will be recorded in the laboratory notebook.

All changes made to the EDS configuration will be noted in the laboratory notebook. Regression testing may need to be performed since the configuration change may change system performance. The Test Manager will determine if regression testing is necessary to ensure the integrity of the collected data.

3.3.4 System Calibration and Threshold Verification.

If applicable to the EDS, system calibration and threshold verification will be conducted periodically during DT&E. These tasks will be performed by the test team and/or vendor personnel according to system operating procedures at the start of every test period and at vendor recommended time intervals thereafter.

3.3.5 Pretest Briefing.

A pretest briefing will be conducted by the Test Manager prior to each test session (e.g. daily). The pretest briefing will be attended by personnel involved with the test effort. This includes assigned test team personnel, vendor representatives, and designated observers. During the pretest briefing, the Test Manager will:

A. Perform detailed safety briefing.
B. Review the objectives of the particular test session
C. Assign specific responsibilities to test personnel.
D. Review system configuration.
E. Review the applicable sections of the Detailed Bulk EDS DT&E Plan.
F. Review the results of relevant tests.
G. Identify and review any existing or expected problems.
H. Review test article set.
I. Review test equipment configuration.

3.3.6 Data Collection.

The Test Manager will assign test team and/or vendor personnel with specific data collection responsibilities prior to execution of each test category. Data collection assignments will be made at the pretest briefing. Designated members will record data in the laboratory notebook, on data collection forms, and/or electronically using data collection equipment, in accordance with the Detailed Bulk EDS DT&E Plan or FAA approved vendor test plans. (The Detailed Bulk EDS DT&E Plan is discussed in Paragraph 4.4.1.)

3.3.7 Problem Reporting and Tracking.

Any test anomaly or equipment problem that occurs during EDS DT&E will be documented. Specific details of the event will be recorded by the test team member who observed the event and validated by the Test Manager immediately following the occurrence (if necessary). The Test Manager will decide if the use of problem reports is appropriate for the test effort. Generally, this documentation is not necessary for the evaluation of systems at early stages of development. The information will be recorded either in the laboratory notebook or on the DT&E Problem Report (DPR) form and will include:

A. Description of observation.
B. Originator name.
C. Date/time occurrence was observed.
D. EDS model number, serial number, software version.
E. Test number and step.
F. Test impact and regression testing recommendation.

At the conclusion of each DT&E test session, all system failure information will be compiled by the test team. The failure data may be used to give an indication of system reliability and potential areas that require further development.

3.4 General Evaluation Approach.

The general evaluation approach includes standard practices for the evaluation of data collected during EDS developmental testing. This section describes evaluation criteria, areas of evaluation, and data reduction and analysis techniques associated with bulk EDS DT&E.

3.4.1 Evaluation Criteria.

EDS evaluation criteria will be defined either in the contract specification or by the Project Manager prior to the start of testing. There may be acceptance criteria which an EDS must meet before being accepted by the Project Manager as a contract deliverable. Data collected from developmental testing performed on EDS's which have reached a mature stage of development (i.e. early production models and/or EDS's approaching contractual acceptance testing) may also be compared to the mandatory certification requirements defined in the applicable certification criteria. EDS certification criteria for checked baggage are contained in Reference D. Additional criteria will be developed for carry-on baggage and electrical items.

3.4.2 Areas of Evaluation.

DT&E areas of evaluation will be dependent on the test objective determined during the planning stage. The following paragraphs describe common EDS DT&E evaluation areas for systems that have reached a mature stage of development. The Project Manager will determine the applicability of these areas of evaluation for each test effort.

By far the most common areas of evaluation involve estimation of probability of detection (p_d) and probability of false alarm (p_{fa}). These are generally approximated by the detection rate and false alarm rate, respectively. Note that detection and false alarm rates are dimensionless and do not refer to rates in terms of time, whereas throughput rate does. These quantities are generally defined by the following ratios:

detection rate:	# alarms ÷ # total threats
false alarm rate:	# alarms ÷ # total non-threats
throughput rate:	# total test articles ÷ processing time (bags/hour, for example).

3.4.2.1 Detection Testing. The objective of detection testing is to determine the rate at which the EDS (operating under a given set of system parameters) can detect different types, quantities, and configurations of explosive materials. For EDS's which identify the location of the explosive in the test article, testing may encompass evaluation of the accuracy of the location identification (true positive determination). The measured detection rate will have a statistical error which is dependent upon the number of independent trials and the required confidence level determined during the planning phase. Data confidence interval calculations are described in Paragraph 3.4.3.1.

3.4.2.2 False Alarm Testing. The objective of false alarm testing is to determine the rate at which the EDS (configured with a given set of system parameters) alarms on non-threat test

articles. False alarm testing may also be conducted to determine what innocuous items and materials cause the EDS to alarm. The measured false alarm rate will have a statistical error which is dependent upon the number of independent trials and the required confidence level determined during the planning phase as described in Paragraph 3.4.3.1.

3.4.2.3 Parametric Testing. The objective of parametric testing is to determine the effect that EDS adjustable hardware or software parameter settings have on EDS performance. Through parametric testing, parameter settings may be determined which optimize certain performance aspects of the EDS. For systems without data collection capabilities or when data collection is not feasible, detection and/or false alarm rate data will be collected for a given set of test articles, at different parameter settings. For systems that have data collection capabilities, raw data may be collected at one set of parameter settings, and detection and false alarm rates can be calculated for other settings using vendor and/or FAA developed software.

In general, DT&E will establish the "operating envelope" of a given EDS by assessing detection probability as a function of false alarm probability. This envelope may be a function of a single system threshold or may depend on highly complex image processing or other data handling procedures. Receiver Operating Characteristic (ROC) curves (alternatively referred to as "p_d/p_{fa} curves") are commonly used to display the relationship between system parameter settings and system detection and false alarm performance. ROC curves are discussed in Paragraph 3.4.3.2. These types of results may be useful in determining the applicability of the particular device or technology to a larger system comprised of several devices. For example, a device with very high detection probability and throughput rate, but high false alarm rate, may serve as an effective "prescreener" to a second device which might have much lower throughput and false alarm rates.

3.4.2.4 Throughput Rate Testing. The objective of throughput rate testing is to determine the rate at which the EDS can process *non-threat* test articles. The elapsed time and number of test articles processed will be recorded during throughput rate testing. If the elapsed time measurement does not include time required for alarm resolution, fault resolution, or system downtime this will lead to a measurement of *maximum* throughput rate. As an EDS is developed further, it may be valuable to measure *average* throughput as well, which would take into account delays due to system inavailability or other factors. Average throughput is particularly important when considering actual operation in an airport.

3.4.2.5 Hazardous Emissions Monitoring. The objective of hazardous emissions monitoring is to ensure that there are acceptable emission levels at critical areas around the EDS. The EDS manufacturer will be required to supply hazardous emissions test data prior to shipping the system to the FAA. The EDS emissions will be checked at the test site prior to test commencement. Locations that are commonly occupied by test personnel during the testing process will be periodically checked with monitoring devices. In addition, all test personnel will wear personal safety equipment as required in applicable Standard Operating Procedures (SOP).

3.4.3. Blind Testing

Testing may also be conducted without test personnel knowledge of which test articles are threat articles and/or what explosive types, amounts, and configurations are being tested. Such testing is termed "blind" testing. In single-blind testing, vendor personnel, or EDS operators, may not have prior knowledge of threat content but FAA data recorders or the Test Manager may. In double-blind testing, FAA test personnel and the Project Manager, in addition, may have no prior knowledge of threat content. However, threat articles likely are prepared with guidance provided by the Project or Test Manager.

3.4.4. Data Reduction and Analysis.

The test team and/or vendor will perform all data reduction and analysis. The data reduction and analysis methods will be defined in the Detailed Bulk EDS DT&E Plan. When conducting baseline testing, identical analysis techniques will be employed for each EDS tested to provide consistent, unbiased results, as specified in the Baseline Test Protocol. The calculation of detection, false alarm, and throughput rates are straightforward "x out of y" calculations. The following paragraphs discuss analysis methods which may be useful in interpreting test results.

Note: Any references to the EDS detection performance with respect to specific quantities of explosives must be classified in accordance with Reference J. Any reference to EDS detection performance with respect to the EDS certification criteria must be classified in accordance with Reference D.

3.4.4.1 Data Confidence Intervals. Data confidence intervals should be calculated during data analysis to determine the statistical error of measured detection and false alarm rates. Data confidence intervals are dependent on the chosen confidence level, number of independent trials, and measured detection or false alarm rate. The confidence interval C is given by:

$$C = \pm z_c \sqrt{p(1-p)/n}$$

where z_c is the confidence coefficient, p is the measured detection or false alarm rate, and n is the number of independent trials. As can be seen from the equation, an increase in the number of trials results in a smaller confidence interval which allows us to more precisely estimate the actual EDS detection and false alarm probabilities. The confidence coefficients are directly related to the chosen confidence levels.

Two-Sided Test. In a two-sided test, one is interested in the possible range of values on either side of the measured proportion (e.g., when one is interested in worst case and best case detection and false alarm rates, or when comparing performance of different EDS's).

One-Sided Test. In a one-sided test, one is only interested in a range of possible values on one side of the measured proportion (e.g., when comparing a measured detection rate against a specific criterion). In detection testing, one may only be interested in the lower confidence

limit of the measured rate, and in false alarm testing, one may only be interested in the upper confidence limit of the measured rate.

The coefficients z_c for several confidence levels are shown in Table 1 for one-sided and two-sided tests.

Table 1. Confidence Coefficients

Confidence Level	60%	70%	80%	90%	95%	99%
z_c (two-sided)	0.842	1.036	1.282	1.645	1.960	2.576
z_c (one-sided)	0.253	0.524	0.842	1.282	1.645	2.326

Example. Suppose that 100 trials are run, and a detection rate of 75% is measured. The two-sided confidence interval for the above example is easily calculated to be ±0.071, which means that one is 90% confident that the actual detection rate is between 67.9% and 82.1%. At the same level of confidence, the one-sided confidence interval is ±0.055, which means that one is 90% confident that the actual detection rate is between 67.9% and 75%, or between 75% and 80.5% (since the interval is symmetric about the actual mean value).

The confidence intervals are the same for a measured proportion p and for $1\text{-}p$. Consequently the confidence intervals for a measured false alarm rate of 25% are the same as those for a measured detection rate of 75%. Therefore, at the 90% confidence level for a two-sided test, the actual false alarm rate is between 17.9% and 32.1%.

3.4.4.2 Receiver Operating Characteristic (ROC) Curves. ROC curves show the relationship between probability of detection, p_d, and probability of false alarm, p_{fa}, for various EDS parameters, such as threshold settings.[1] An ROC curve may be generated by setting the EDS threshold at a certain value, running a number of threat and non-threat test articles through the system, and calculating the detection and false alarm rates. The test is then repeated at different threshold settings. Each test is represented by one point on the curve. This method is likely to become impractical if many p_d/p_{fa} values are desired.

If the EDS can record the relevant raw data for off-line computer analysis, then the test only needs to be performed once. ROC curves can then be generated using vendor and/or FAA developed software.

Some EDS's have many adjustable hardware and software parameters which affect the detection and false alarm performance of the system. For these systems, coordination with the EDS vendor may be necessary in order to determine which parameters to vary with respect to the explosive type and quantity under test. For example, a range of values of p_{fa} can be specified in advance, and software optimized (by the vendor) to maximize p_d.

[1] H. Urkowitz, Signal Theory and Random Processes, Artech House, Inc., Dedham, MA, 1983, page 522.

Two fictitious ROC curves are shown in Figure 3. This graph shows potential system response for two masses of the same type of explosive. Note the increasing tradeoff between probability of detection and probability of false alarm at the top of the curves. (For the Mass=1.0 explosive, an increase in p_d from 90% to 92% results in a subsequent increase in p_{fa} from 30% to 50%.)

NOTE: Vendor assistance with test planning, performance and/or analysis should be in conformance with applicable contract requirements or other agreements.

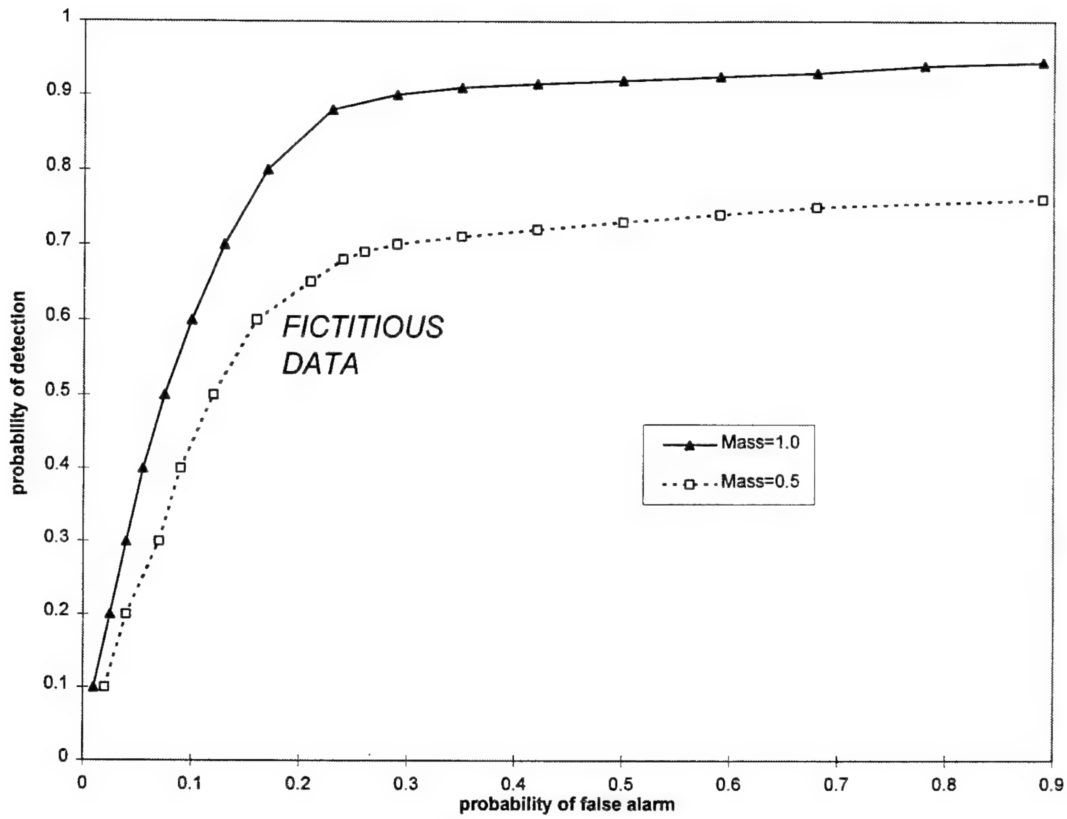


Figure 3. Fictitious ROC Curves.

4. TEST MANAGEMENT

This section defines the test management approach to be taken throughout DT&E, and includes a description of training requirements, safety requirements, test schedule, documentation requirements, test control, and roles and responsibilities.

4.1 Training.

Where agreements between the FAA and EDS vendors exist, the vendor will provide operational training to the test team prior to FAA conducted DT&E. The level of training will be sufficient to enable the test team to independently set up, initialize, and operate the EDS during DT&E. For vendor conducted DT&E, the vendor may be required to provide familiarization training to the test team. This level of training will be sufficient to enable the test team to operate the system.

4.2 Safety.

All DT&E conducted at the FAA Technical Center Aviation Security Laboratory will be in accordance with safety plans and procedures contained in References E through H. DT&E conducted at vendor facilities will be in accordance with vendor safety plans and procedures. In the event that the vendor does not have established safety plans and procedures, applicable FAA safety plans and procedures will be utilized.

4.3 Test Schedule.

The Project and Test Managers will be responsible for developing a schedule for all DT&E events. The schedule will include laboratory preparation, documentation preparation, test conduct, and data analysis activities.

4.4 Documentation Requirements.

4.4.1 Detailed Bulk EDS DT&E Plan.

In order to facilitate the testing of a specific EDS and accommodate unique test objectives and scenarios, a detailed test plan will be required for each test effort. The plan will be reviewed by the Program/Project Manager, Test Director, Safety Officer, Lab Manager, and the EDS vendor, prior to test commencement. FAA developed bulk EDS DT&E plans will contain the following information, at a minimum:

- A. Purpose
- B. Scope
- C. Reference Documents
- D. System Description
- E. Test Description
    - E.1. Test Objectives

E.2. Test Approach (e.g., data collection methods, data confidence levels, configuration management, system calibration, safety precautions, specific procedures, security procedures)
E.3. Test Resource Requirements (e.g. personnel, equipment, facility, proposed schedule, test articles)
E.4. Data Reduction and Analysis Methodology

Vendor-developed test plans will also contain the above information, consistent with contract requirements or other applicable agreements. If vendor-developed test plans do not contain all the required information, the Project Manager will include the information in the test plan.

4.4.2 Laboratory Notebook

The Test Manager may choose to utilize a laboratory notebook for the recording of test data and observations, test problems, and data analysis. At a minimum, a laboratory notebook is required for the recording of test observations, configuration data, and general notes. The Test Manager will have the option of recording test data and test problems in the laboratory notebook, on custom data collection forms, or on the standard forms included in Appendix A. The laboratory notebook will remain an FAA internal unpublished document used in preparing the test report (see below) and retained for archival purposes.

4.4.3 Bulk EDS DT&E Report.

The Bulk EDS DT&E Report will contain a detailed analysis of the developmental testing results, and a summary of any problems encountered during the test. The report will contain all the information necessary to evaluate the system with respect to test objectives and criteria. This report will be prepared by the Project Manager based on results from data reduction and analysis reported by the Test Manager. Even though the Project Manager and Test Manager may be the same person, the Project Manager is expected to be the person most familiar with the system under test and in a position to formulate conclusions and recommendations based on the results of the test. Each DT&E test report will be reviewed by the Program Lead.

All bulk EDS DT&E reports will contain the following information, at a minimum:

A. Executive Summary
B. Test Purpose
C. Reference Documents
D. System Description/ Configuration
E. Test Date and Location
F. Test Participants
G. Test Description
    G.1. Test Objectives
    G.2. Test Conduct (e.g. data collection, reference to specific procedures, procedural deviations, test article set)
    G.3. Data Reduction and Analysis Methodology
H. Detailed Test Results

I. Test Anomalies
J. Conclusions/Recommendations
K. Appendix (e.g. test article documentation)

The vendor will prepare test reports for vendor conducted DT&E, as required by contract or other established agreements with the FAA. These reports will be reviewed and approved by the Project Manager. An independent analysis verification may also be conducted by the Project Manager.

4.4.4 Guidance Documentation.

In addition to this plan, other documentation has been developed to be used as guidance during bulk EDS DT&E. This documentation is described in the following paragraphs.

4.4.4.1 DT&E Test Article Selection Plan. The DT&E Test Article Selection Plan provides guidance for the development of the standard test article set and custom test article sets. The plan will provide test article selection criteria and methodology, and explosive parameters to be used in test article set development.

Since the standard test article set may be assembled and disassembled for each DT&E effort, the plan will provide detailed guidance on the construction of the standard set (e.g. explosive parameters, configurations, and locations in the test articles). Since the custom test article set will vary from test to test, the plan will provide general guidance on the construction of the custom set (e.g. data confidence intervals, explosive location determination via random number generation).

4.4.4.2 Baseline Test Protocol. The Baseline Test Protocol provides a specific set of procedures for the conduct of baseline testing utilizing the standard test article set. Baseline testing will provide system performance data with sufficient statistical precision to allow for meaningful comparisons of systems and technologies.

4.5 Test Control.

Control of test activities and documentation of results is required throughout all phases of DT&E. To accomplish this, test control documentation will be utilized. This documentation will include a configuration log, test data, and DT&E problem reports. The Test Manager will determine what amount of control documentation is appropriate for each DT&E effort.

4.5.1 Test Log.

The laboratory notebook will be used to document test activities and control EDS configuration throughout DT&E. The laboratory notebook will contain summaries of all aspects of the test, such as start and stop of test sequences, progress of achieving test objectives, problems which occur and other relevant information. EDS configuration data will be recorded in the event any system baseline changes occur during testing. Entries into the notebook will be validated by the Test Manager, when appropriate.

4.5.2 Test Article Documentation.

Information on all test articles will be recorded on standard forms from Appendix A or other appropriate documents. Information on explosive sample weights, types, locations, configurations, orientations, etc. will be recorded. Relevant additional information on luggage or electronics test articles such as item number, type, size, weight, etc.. Information on interference samples incorporated into test articles will be recorded.

4.5.3 Test Data.

The test data defined in Paragraph 3.3.6 may be collected throughout DT&E and will be retained for subsequent analysis. Test data will be recorded either in the laboratory notebook, on custom data collection forms, on the standard forms included in Appendix A, or into a computer file or database. Proper recording of these data is the responsibility of the Test Manager. Any EDS proprietary information obtained during DT&E will be used for test and analysis purposes only.

4.5.4 DT&E Problem Report (DPR).

DPR's may be written during testing to document test anomalies and equipment failures. Specific details will be recorded on the DPR form and will contain, at a minimum, the information described in Paragraph 3.3.7. All entries will be recorded by the test team member who observed the event and validated by the Test Manager immediately following the occurrence. In general, a DPR will be written if the observation is related to equipment failure or may affect test results. All problem reports will be addressed in the DT&E test report. This will include a description of the problem, disposition, and the results of any retests, if applicable.

4.6 Roles And Responsibilities

Bulk Explosives Detection Program Lead. The Program Lead is responsible for overall direction of bulk EDS DT&E efforts including:

A. Develop standard set of test articles for bulk EDS DT&E.
B. Designate Project Manager and Test Manager for DT&E efforts, as required.
C. Coordinate with Test Director, as required.
D. Review and approve each Detailed Bulk EDS DT&E Plan.
E. Review and approve each Bulk EDS DT&E Report.

Project Manager. The Project Manager will be responsible for overall test management including:

A. Determine test objectives, evaluation criteria, confidence levels, and schedule.
B. Schedule all required equipment and resources.
C. Develop custom test article set, as required.
D. Develop the Detailed Bulk EDS DT&E Plan.

E. Prepare the Bulk EDS DT&E report.
F. Review vendor test plans and procedures, and reports.
G. Conduct independent analysis verification of vendor test results.
H. Determine applicability of the contents of this plan for each test effort.

Test Manager.

A. Ensure the EDS is properly installed and operational prior to test conduct.
B. Ensure the EDS meets all safety requirements prior to test conduct.
C. Review and conduct test in accordance with Detailed Bulk EDS DT&E Plan.
D. Conduct pretest briefings.
E. Coordinate activities of all test personnel during test conduct.
F. Ensure compliance with test procedures.
G. Oversee collection and reduction of test data.
H. Approve system configuration changes.
I. Perform data reduction and analysis.
J. Provide results of analysis to Project Manager.

EDS Vendor. In accordance with contractual requirements or other agreements, the vendor will:

A. Develop test plan.
B. Install and maintain equipment.
C. Provide training to test team personnel.
D. Perform test operations.
E. Prepare DT&E test report.
F. Provide hazardous emissions monitoring equipment.

Test Director.

A. Support development of a standard set of test articles for EDS DT&E.
B. Support development of a custom test article set, as required.
C. Coordinate with Program/Project Manager, as requested.
D. Review the Detailed Bulk EDS DT&E Plan, as requested.
E. Support execution of DT&E efforts as requested by the Project Manager.

Test Team.

A. Operate EDS during DT&E.
B. Collect and record test data.
C. Record daily test logs and other documentation.
D. Assist Test Manager as required.

Explosives Technician(s). The explosives technicians will be FAA and/or vendor personnel and will:

A. Verify explosive sample availability.
B. Obtain explosive samples, as required.
C. Prepare all explosive samples and test articles containing explosives.
D. Perform explosive and test article handling, as required.
E. Ensure compliance with safe handling and storage procedures.

Safety Officer.

A. Validate that the EDS meets all safety requirements prior to test conduct.
B. Ensure compliance with all applicable health and safety standard operating procedures (References E-H) during test conduct.
C. Conduct all required safety briefings.
D. Review and approve Detailed Bulk EDS DT&E Plan.

APPENDIX A. BULK EDS DT&E STANDARD FORMS

Standardized forms which can be used in bulk developmental test and evaluation are presented in the following order:

1. Data Collection Form: Detection Testing
2. Data Collection Form: False Alarm Testing
3. Data Collection Form: Throughput Testing
4. Data Collection Form: Hazardous Emissions
5. Threat Article Definition Sheet for Checked/Carry-on Baggage: Bulk Explosive Locations
6. Threat Article Definition Sheet for Checked/Carry-on Baggage: Sheet Explosive Locations
7. Threat Article Definition Sheet for Electrical Items
8. DPR: Bulk EDS DT&E Problem Report

(this page intentionally left blank)

DATA COLLECTION FORM
DETECTION TESTING

DETECTION TESTING					
TEST ARTICLE #	EDS RESPONSE	COMMENTS	TEST ARTICLE #	EDS RESPONSE	COMMENTS

EDS TESTED:

____

TEST ENGINEER: ______________________ DATE:

____

TEST MANAGER: ______________________ PAGE _____ OF _____

DATA COLLECTION FORM
FALSE ALARM TESTING

FALSE ALARM TESTING					
TEST ARTICLE #	EDS RESPONSE	COMMENTS	TEST ARTICLE #	EDS RESPONSE	COMMENTS

EDS TESTED: ______________________________

TEST ENGINEER: ____________________ DATE: ____________

TEST MANAGER: PAGE OF

DATA COLLECTION FORM
THROUGHPUT TESTING

THROUGHPUT TESTING		
TIME ELAPSED	NUMBER OF TEST ARTICLES PROCESSED	COMMENTS

EDS TESTED:

____

TEST ENGINEER: ______________________ DATE: ________________

TEST MANAGER: ______________________ PAGE _____ OF _____

DATA COLLECTION FORM
HAZARDOUS EMISSIONS

EDS HAZARDOUS EMISSIONS				
MONITORING LOCATION	TIME	RADIATION LEVEL (mR/hr) small bag in EDS	large bag in EDS	COMMENTS
A				
B				
C				
A				
B				
C				

SKETCH DIAGRAM OF EDS BELOW

EDS TESTED:

____

MANUFACTURER / MODEL NO. / SERIAL NO. / CAL. EXP. DATE

TEST EQUIPMENT:

____

PREPARED BY: ______________________ DATE: ________________

(Test Engineer)

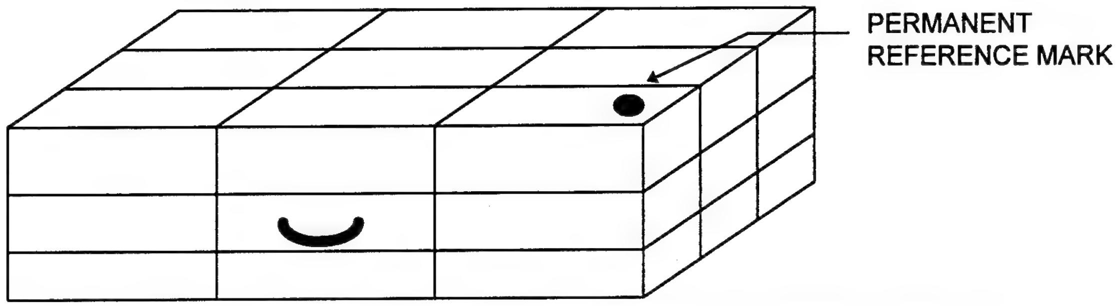
VALIDATED BY: ______________________

(Safety Officer)

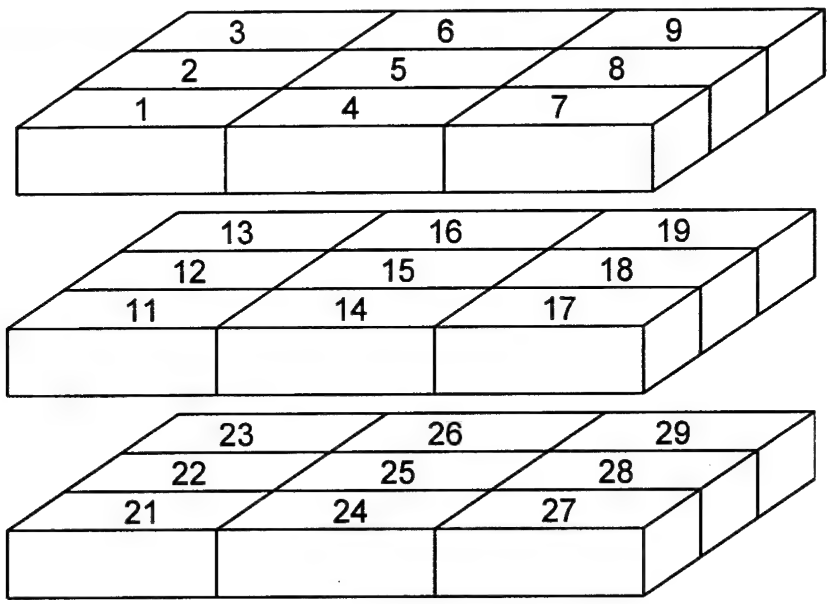
THREAT ARTICLE DEFINITION SHEET
for CHECKED/CARRY-ON BAGGAGE

THREAT ARTICLE #______

BULK EXPLOSIVE LOCATIONS



MARK DIAGRAM TO SHOW LOCATION OF EXPLOSIVE IN THE TEST ARTICLE



EXPLOSIVE

Type: ____

Mass: ____________________ Dimensions (L×W×H): ________________

Memo: ____

____

____

____

____

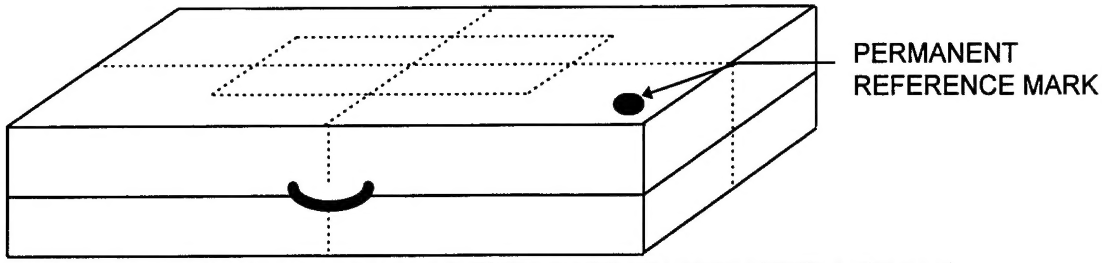
PREPARED BY: ______________________ DATE: ____________
(Explosive Specialist)

VALIDATED BY: ______________________ DATE: ____________
(QA Representative)

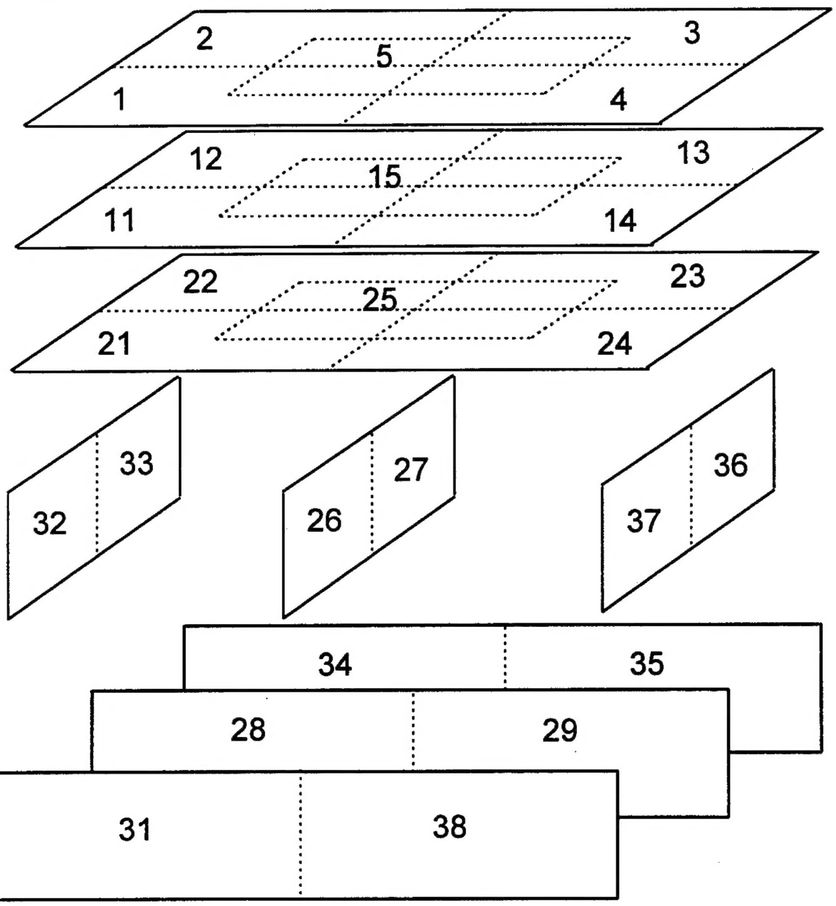
THREAT ARTICLE DEFINITION SHEET
for CHECKED/CARRY-ON BAGGAGE

THREAT ARTICLE # ________

SHEET EXPLOSIVE LOCATIONS



MARK DIAGRAM TO SHOW LOCATION OF EXPLOSIVE IN THE TEST ARTICLE



EXPLOSIVE

Type: ______________________________ Thickness: ________________

Mass: ____________________ Dimensions (L×W): __________________

Memo: ____

____

PREPARED BY: ________________________ DATE: ________________
(Explosive Specialist)

VALIDATED BY: ________________________ DATE: ________________
(QA Representative)

THREAT ARTICLE DEFINITION SHEET
for ELECTRICAL ITEMS

THREAT ARTICLE #______

BULK or SHEET EXPLOSIVE LOCATIONS

SKETCH ELECTRICAL ITEM BELOW
MARK DIAGRAM TO SHOW LOCATION OF EXPLOSIVE IN THE TEST ARTICLE

EXPLOSIVE

Type:

______________________________ Thickness: _________________________

Mass: ________________________ Dimensions (L×W×H):

____

Memo:

____

____

____

PREPARED BY: ___________________________ DATE: ___________________
(Explosive Specialist)

VALIDATED BY: ___________________________ DATE: ___________________
(QA Representative)

DPR

BULK EDS DT&E PROBLEM REPORT

1. DPR LOG DATE / /	2. DPR NUMBER DPR________

ORIGINATOR/SYSTEM INFORMATION:

3. ORIGINATOR NAME	4. DATE OBSERVED / /
5. SYSTEM MODEL/NUMBER	6. SYSTEM SERIAL NUMBER
7. SOFTWARE VERSION	8. TEST ARTICLE NUMBER (IF APPLICABLE)
9. TEST NUMBER	10. TEST STEP

PROBLEM INFORMATION:

11. SHORT TITLE (ONLY WRITE IN THE SPACES PROVIDED)

| | | | | | | | | | | | | | | | | | | | | | | | | | | | |

12. DESCRIPTION OF PROBLEM (UNCLASSIFIED INFORMATION ONLY)

ATTACHMENT [ ]

13. TEST IMPACT/REGRESSION TEST REQUIREMENTS

14. ASSIGNED PRIORITY (CIRCLE ONE) I II III	

APPROVED BY________________________
(TEST MANAGER)

DPR1.FRM
REVISION 2
12-28-94

DISTRIBUTION LIST

ORGANIZATION	ADDRESS	NAME	NO. COPIES
AAR-500	Aviation Security Research and Development Division (Library)	Bethany Prehodka (DCS)	2
ACT-5	FAA Technical Center Office of Public Affairs		1
ACP-400	Federal Aviation Administration Office of Civil Aviation Security Policy and Planning 800 Independence Ave. S.W. Washington, DC 20591	Keith Goll	1
ACS-20	Federal Aviation Administration Office of Civil Aviation Security 800 Independence Ave. S.W. Washington, DC 20591	Lyle Malotky	1
M493.2	DOT Library FB10A, M457, Room 930 800 Independence Ave., S.W. Washington, DC 20591		4
DCS, Inc.	DCS, Inc. 2511 Fire Road Suite A5 Absecon, NJ 08201		2